Technological Forecasting & Social Change 95 (2015) 96–108

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Detecting tension in online communities with computational Twitter analysis

Pete Burnap [a,*], Omer F. Rana [a], Nick Avis [a], Matthew Williams [b], William Housley [b], Adam Edwards [b], Jeffrey Morgan [b], Luke Sloan [b]

[a] Cardiff School of Computer Science & Informatics, Cardiff University, Cardiff, UK
[b] Cardiff School of Social Sciences, Cardiff University, Cardiff, UK

ARTICLE INFO

Article history:
Received 12 October 2012
Received in revised form 18 April 2013
Accepted 19 April 2013
Available online 11 May 2013

Keywords:
Opinion mining
Sentiment analysis
Text mining
Social media analysis
Machine learning
Conversation analysis
Membership categorization analysis

ABSTRACT

The growing number of people using social media to communicate with others and document their personal opinion and action is creating a significant stream of data that provides the opportunity for social scientists to conduct online forms of research, providing an insight into online social formations. This paper investigates the possibility of forecasting spikes in social tension – defined by the UK police service as "any incident that would tend to show that the normal relationship between individuals or groups has seriously deteriorated" – through social media. A number of different computational methods were trialed to detect spikes in tension using a human coded sample of data collected from Twitter, relating to an accusation of racial abuse during a Premier League football match. Conversation analysis combined with syntactic and lexicon-based text mining rules; sentiment analysis; and machine learning methods was tested as a possible approach. Results indicate that a combination of conversation analysis methods and text mining outperforms a number of machine learning approaches and a sentiment analysis tool at classifying tension levels in individual tweets.

1. Introduction

Social networking technologies, such as those provided by online social networking sites such as Facebook (where users have an online social "friendship" relationship), and microblogging websites such as Twitter (where the online social relationship can be uni-directional or reciprocal), provide people with unprecedented interactive opportunities. These services also support the dissemination, discussion and often distortion of information within and between online and offline 'communities' at speeds hitherto not possible. The blurring of the boundary between offline and online discourse requires us to rethink the definition of human-to-human interaction. In this paper we study this new form of interaction by examining how events that occur offline have the potential to affect terrestrial societal cohesion and order, and how they are reflected in discourse conducted through online social media. We argue that offline events can trigger 'online tension' that can be measured via a series of indices or metrics. These metrics can then be used to monitor peaks in tension in social media, affording a form of 'anticipatory governance'.

Tension is defined by the Police Service in the United Kingdom as "any incident that would tend to show that the normal relationship between individuals or groups has seriously deteriorated and is likely to escalate to wider groups other than those involved" [1]. It is therefore important to study the possibility that collecting, analyzing and visualizing self reported information posted to online social networking sites could provide an insight into online tension through the mining and analysis of public opinion.

Opinion mining of social media data has received a lot of research attention in recent years (e.g. [2–5]), largely due to the increasing use of social networking technologies that enable citizens to self-report their opinions as frequently as they wish on a wide range of topics, and the availability of programmatic access to such data through application programming interfaces (APIs). The method of opinion

* Corresponding author. Tel.: +44 2920874000.
E-mail address: p.burnap@cs.cardiff.ac.uk (P. Burnap).

0040-1625/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.techfore.2013.04.013

mining generally requires the identification of: an entity to which the opinion is focused on (e.g. a person); attributes of the entity (e.g. the person's political perspective); views, attitudes or feelings towards the entity and its attributes (commonly defined as sentiment); an opinion holder; and a time at which the sentiment was expressed [6].

One of the key research challenges for textual opinion mining has been to classify sentiment expressed by the opinion holder, or a collection of opinion holders (e.g. comments posted in response to a news story), using a pre-defined set of classes (e.g. positive, negative or neutral). To achieve this, it is typical to try and identify features within the text that are useful for deciding which class it belongs to. These features may be positive or negative words (e.g. good, excellent, bad, and awful) or descriptive words in the context of the entity (e.g. this person is wrong about this = negative). Features may also be syntactic, such as verbs or nouns, or words that belong to particular lexicons (e.g. common swear words). In social media there are idiosyncratic features such as hashtags (the # symbol) and emoticons, which can also be used to inform classification decisions.

The Cardiff Online Social Media ObServatory (COSMOS) is a Web Observatory platform developed to support researchers interested in collecting, analyzing and visualizing publicly accessible digital data feeds [7,36-37]. One specific objective of the COSMOS platform is to develop computational tools to mine opinion from social media data and enable the detection of tension indicators in online communities, visualizing them such that spikes in tension can be detected. A spike exists where tension can be measured at a single point in time as significantly higher than it has been previously — an anomaly in a timeline.

Using the definition of tension as defined in [1] – "any incident that would tend to show that the normal relationship between individuals or groups has seriously deteriorated and is likely to escalate to wider groups other than those involved" – this study is specifically focused on developing an opinion mining application capable of classifying public posts, published through the social media microblogging site Twitter, with respect to the level of tension expressed. In terrestrial communities, such as particular neighborhoods and geographic regions, tension indicators are often obvious and apparent to the naked eye. Broken windows, graffiti, banners and posters are all potential indicators of tension within a locality. Tension indicators in online communities are not so obvious, but Twitter is a very open self-reporting platform that is underpinned by an online social network and has recently been shown to host discussions around sensitive topics such as riots [8], political campaigns [4], and the Tunisian and Egyptian revolutions [9,10]. Hence, messages posted to Twitter – 'tweets' – are the device through which it is proposed that tension can be detected.

The premise of the study from a forecasting perspective is that it will be possible to detect a rise in tension levels over time in relation to an observed event. Such a rise would thus be indicative of growing online tension, and allows those observing online tension to react accordingly. With the growing use of social media for expressing opinion, and the relative uncertainty around its interpretation, it is expected that a visualization of tension classified at a number of levels, with each level being plotted over time, would facilitate the observation process. From this, increases in the plotted tension levels – particularly the higher levels – are visible and therefore could be used as a metric for use in forecasting disruptive social phenomena during, and following, a known event.

In this paper we present the results of a tension classification study using a number of computational approaches. In Section 2 we provide a summary of the state-of-the-art in computational sentiment analysis and its application to study social phenomena. In Section 3 we discuss data collection and annotation along with the various methods used in the study and how they build on existing best practice. In Section 4 we discuss the results of applying different computational techniques to the problem of tension classification. In particular, the aim of this task was to test the performance of a number of content analysis and machine learning approaches with the intended outcome of identifying the best performing methods for classifying tension on a number of levels. We were motivated by the assumption that no single method would perform best at every level of tension. Finally, we discuss the benefits, relevance, and limitations of the outcomes in Section 5.

2. Background

Qualitative and quantitative research have been re-framed in the light of the rise of Web 2.0. Rathi and Given propose a framework for research in Web 2.0 environments (Research 2.0), which considers the web as a research platform, harnessing the power of crowds, and creating research databases from expansive online content that they refer to as perpetual data and coin the term Data 2.0 [11]. Data can be obtained from many sources such as personal blogs, wikis, microblogs posted on social networking sites (e.g. Facebook and MySpace status updates and Twitter tweets) and RSS (really simple syndication) feeds from sources such as news reporting websites [12].

These data, combined with emotive opinion mining methods, have been used in several recent sociological and economic academic studies. Computational statistical data mining methods have been used to investigate the demographics and 'friending' behavior on MySpace [13,14]. Tailored social media data harvesting methods, combined with human interpretation, have been used to investigate the change in people's ideology concerning the privacy of childbirth [15]. In [5] the authors conduct a large scale study of online social networks to extract different types of mood using sentiment analysis and the Profile of Mood States (POMS) psychological mood scale, and relate changes in mood to real-world socio-economical events. Building on this, psychological "well being" states have been monitored over time to show that online social networking reflects the assortive nature exhibited offline [2]. Sentiment analysis has been used to determine emotional differences between genders on MySpace [13,14], and study levels of positive and negative sentiments in Facebook [16] and Twitter comments [2,3]. Sentiment analysis has also been used to 'predict' election outcomes [4], and it was demonstrated that sentiment relating to new movie releases, combined with tweet frequency, was more accurate at predicting revenue than the Hollywood stock market [17]. For forecasting purposes sentiment analysis appears to offer a significant insight into online communication.

Three common approaches to sentiment analysis are text-based machine learning, lexicon-based methods and linguistic